DOCUMENT RESUME

ED 173 354 TM 008 279

AUTHOR Catts, Ralph
TITLE Q. How Many Options Should a Multiple-Choice Question Have? (a) 2. (b) 3. (c) 4. At-a-glance Research Report.
INSTITUTION New South Wales Dept. of Education, Sydney (Australia).
PUB DATE [78]
NOTE 11p.

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Answer Sheets; Complexity Level; Foreign Countries; Higher Education; *Multiple Choice Tests; Reaction Time; Response Mode; *Test Construction; *Test Items; *Test Reliability
IDENTIFIERS Australia; *Distractors (Tests); *Test Length

ABSTRACT

The reliability of multiple choice tests--containing different numbers of response options--was investigated for 260 students enrolled in technical college economics courses. Four test forms, constructed from previously used four-option items, were administered, consisting of (1) 60 two-option items--two distractors randomly discarded; (2) 40 three-option items--one distractor randomly discarded; (3) 40 three-option items--the least appropriate distractor discarded; and (4) 30 of the original four-option items. Data from students who didn't complete the test or who omitted many items were disregarded. Results indicated that three-option items were as reliable as four-option items (reliabilities were .6 to .7 and .7, respectively); that higher scores were obtained on tests with fewer options per item; and that tests consisting of four-option item required more time for both test construction and administration. A comparison of tests (2) and (3), both containing three-option items, indicated that test (2) was more reliable, less difficult, and took longer to complete. (GDC)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

TM008 279

ED173354

At-a-glance Research Report

U S DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Ralph Catts

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."

from the

Assessment Research & Development Unit (A.R.D.U.)

N.S.W. Dept. of Technical & Further Education

Q. How many options should a multiple-choice question have?

(a) 2
(b) 3
(c) 4

– Ralph Catts.

ISBN 7240 - 3462 - 5

The data on which this report is based was collected in 1973 with the assistance of the staff of the School of General Studies. The project was funded by the N.S.W. Department of T.A.F.E., and was jointly designed and conducted by

Ralph Straton (formerly at the Uni. of Sydney)
and Ralph Catts (Assessment R & D Unit, TAFE).

A more comprehensive study involving staff from other Schools within the N.S.W. Department of T.A.F.E. is being planned.

Q. How many options should a multiple-choice question have?

(a) 2
(b) 3
(c) 4

Introduction

Every teacher who has tried to write multiple choice items knows that the more options you attempt to write, the harder it becomes. On the other hand, it would seem that the more options you include, the less chance there is of the student guessing the correct answer.

Given this later widespread view, it may come as a surprise to learn that several writers have suggested that tests having items with just 3 options (including the correct answer) are just as "reliable" as tests with items having a greater number of options. The study reported here investigated this claim about 3 option items using a group of TAFE students enrolled in the Diploma Entrance Course.

(The "reliability" of a test refers to how consistently it measures whatever it is supposed to measure. Obviously high reliability is an essential feature of a good test. Reliability coefficients range from a minimum of zero to a maximum of one. Good classroom tests typically have reliability coefficients between 0.6 and 0.8.)

Method

The Students: 260 students enrolled in Day Matriculation economics courses at 8 Sydney suburban technical colleges were the subjects in this study.

The Tests: The traditional "trial examination" was used to provide a realistic setting where motivation to do well would be high. 60 items were selected from previous H.S.C. (Higher School Certificate) exams to form an item pool from which alternate forms of the test paper could be created. All items in the original pool contained 4 options. Items with 2 or 3 options were produced by discarding distractors from the original items.

Test A contained 2 option items (2 distractors randomly discarded)
Test B contained 3 option items (1 distractor randomly discarded)
Test C contained 3 option items (1 distractor selected as least appropriate and discarded)
Test D contained 4 option items (taken directly from item pool)

For reasons relating to previous research in this field, it was important that the total number of options in each test was the same. Therefore the alternate forms of the test were as follows.

	TEST A	TEST B	TEST C	TEST D
Number of items	60	40	40	30
Number of options per item	2	3	3	4
Total number of options	120	120	120	120

As it is also useful to have measures of reliability and completion times for tests with the same number of items, results are reported both for tests with equal total number of options and, after necessary adjustments*, for tests with equal numbers of items.

Test Administration: Students were given the test in their normal classes. Teachers distributed the 4 versions of the test in a random fashion around the class in order to negate the effects of grouping and seating. The responses of students who did not complete the test, or who omitted a high proportion of items were not included in the analysis.

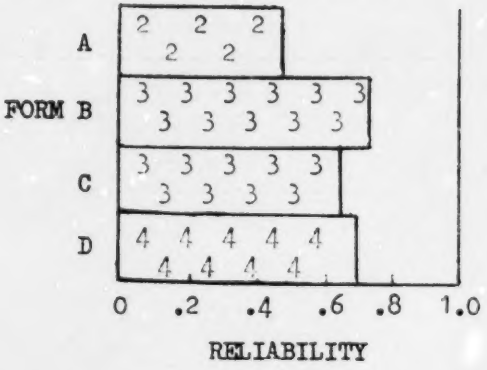
---

* Simple extrapolation was used to estimate completion times, and the Spearman-Brown formula was used to estimate reliabilities of tests with equal numbers of items.

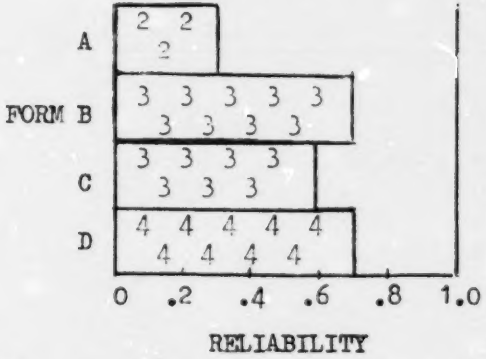
Results

1. TEST RELIABILITY

1. EQUAL NUMBER OF OPTIONS



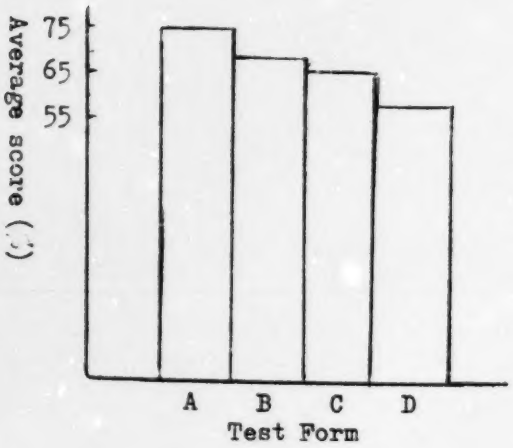
2. EQUAL NUMBER OF ITEMS



Irrespective of whether you consider tests with equal total number of options or with equal numbers of items, 3 and 4 option tests don't differ to any significant extent in their reliabilities, but 2 options per item produce inferior reliabilities.

Note also that a test consisting of 30 items can be as reliable as a test consisting of 60 items.

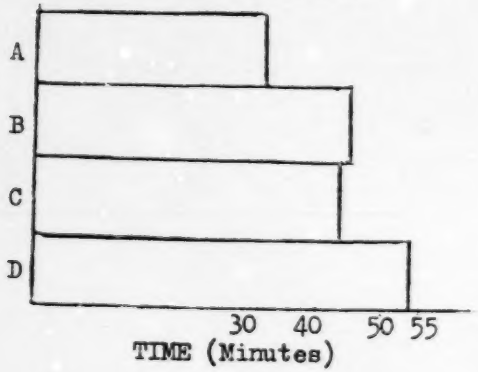
2. TEST DIFFICULTY



As might be expected, the more options per item, the more difficult the test.

3. TIME TAKEN TO COMPLETE TEST

EQUAL NUMBER OF ITEMS



Not only do four options per item tests take longer to construct, but they also take longer for the student to complete.

CONCLUSIONS AND RECOMMENDATIONS

The results of this study agree with earlier research which suggests that three option multiple choice items are as good as four option multiple choice items.

Tests with three options per item appear to be:

- as reliable as tests with four options per item
- take less time for students to complete
- slightly easier than equivalent tests with four options per item.

For those who are concerned with the effects of this last outcome on pass rates, the solutions are easy. Either select items for your test with harder distractors or, adjust the pass mark to compensate for the slightly easier test.

Further studies need to be carried out to find out if these findings hold up with different types of students and with different test content. Remember, these results are for Day Matriculation students studying Economics. Your students, and your content area may produce different results. If you'd like to find out how many options per item you should be using contact the author at ARDU (20982, Ext. 9254) and we will help you.

AT-A-GLANCE RESEARCH REPORTS

The aim of this series is to focus on issues in student assessment by reporting and discussing studies conducted within the N.S.W. Department of TAFE.

The At-a-Glance reports are summaries of research carried out by the Assessment Research and Development Unit. Each report deals with some aspect of:-

- student assessment;
- student selection and placement; or
- evaluations of TAFE programs and policies.